

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 14:56:48 ; Search time 41 Seconds  
(without alignments)  
1951.129 Million cell updates/sec

Title: US-09-831-805A-6

Perfect score: 310

Sequence: 1 MALRRPRLRLCARLPDFLL.....VNYINTDEGDRHKSFEVI 310

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

SPTREMBL\_23.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	209	67.4	310	4	Q9BX67
2	209	67.4	355	4	Q8WML8
3	134	43.2	309	4	Q96FL1
4	32	10.3	310	11	Q9D8B7
5	28	9.0	310	11	Q9D1M9
6	28	9.0	310	11	Q9EPK4
7	13	4.2	64	11	Q8BT59
8	8	2.6	39	7	Q9TNT2
9	8	2.6	67	7	Q8HWG1
10	8	2.6	131	5	Q8I8S1
11	8	2.6	154	11	Q8C4V2
12	8	2.6	231	17	Q8TQ75
13	8	2.6	232	17	Q8ZSR8
14	8	2.6	233	10	Q9LRA9
15	8	2.6	246	5	Q966C0
16	8	2.6	290	7	Q29990

17	8	2.6	294	16	Q98EQ9	Q98eq9 rhizobium l
18	8	2.6	309	5	Q9N4Y9	Q9n4y9 caenorhabdi
19	8	2.6	313	7	Q9MY31	Q9my31 homo sapien
20	8	2.6	322	2	Q9RP17	Q9rp17 desulfitoba
21	8	2.6	336	10	Q9C7E9	Q9c7e9 arabidopsis
22	8	2.6	338	7	Q9MXL5	Q9mxl5 pan troglod
23	8	2.6	348	7	Q9MWJ9	Q9mwj9 gorilla gor
24	8	2.6	348	16	Q9PR46	Q9pr46 ureaplasma
25	8	2.6	363	7	Q95HC2	Q95hc2 homo sapien
26	8	2.6	365	7	Q9XR88	Q9xr88 homo sapien
27	8	2.6	365	7	Q9XR77	Q9xr77 pongo pygma
28	8	2.6	365	7	Q9MXG4	Q9mxg4 pan troglod
29	8	2.6	366	6	Q46684	Q46684 pan troglod
30	8	2.6	366	7	Q8MH00	Q8mh00 pongo pygma
31	8	2.6	366	7	Q9TQ99	Q9tqp9 homo sapien
32	8	2.6	366	7	Q29652	Q29652 homo sapien
33	8	2.6	366	7	Q95463	Q95463 homo sapien
34	8	2.6	366	7	Q8MHH1	Q8mhh1 pongo pygma
35	8	2.6	366	7	Q8MHG8	Q8mhg8 pongo pygma
36	8	2.6	366	7	Q8MHG9	Q8mhg9 pongo pygma
37	8	2.6	366	7	Q19617	Q19617 homo sapien
38	8	2.6	366	7	Q78083	Q78083 homo sapien
39	8	2.6	366	7	Q29865	Q29865 homo sapien
40	8	2.6	366	7	Q95603	Q95603 homo sapien
41	8	2.6	366	7	Q9TPL2	Q9tpl2 pan troglod
42	8	2.6	366	7	Q9MXD2	Q9mxd2 pan troglod
43	8	2.6	449	16	Q8CWZ7	Q8cwz7 streptococc
44	8	2.6	452	16	Q8G5D2	Q8g5d2 bifidobacte
45	8	2.6	492	16	Q8ZIU9	Q8zlu9 salmonella
46	8	2.6	659	16	Q8EK37	Q8ek37 shewanella
47	8	2.6	784	16	Q9CIF4	Q9cif4 lactococcus
48	8	2.6	1290	13	Q9W6E1	Q9w6e1 gallus gall
49	7	2.3	67	4	Q8NI85	Q8ni85 homo sapien
50	7	2.3	85	12	Q9GCX8	Q9gxc8 helicoverpa
51	7	2.3	86	16	Q8ERD1	Q8erd1 oceanobacil
52	7	2.3	92	10	Q8LFP4	Q8lfp4 arabidopsis
53	7	2.3	92	10	Q8GXY9	Q8gxy9 arabidopsis
54	7	2.3	104	9	Q8W6S9	Q8w6s9 bacterioph
55	7	2.3	112	16	Q8CSN1	Q8csn1 staphylococ
56	7	2.3	113	16	Q928J9	Q928j9 listeria in
57	7	2.3	116	17	Q8U3Z7	Q8u3z7 pyrococcus
58	7	2.3	135	2	Q8RJ00	Q8rj00 streptococc
59	7	2.3	144	10	Q8LN01	Q8ln01 oryza sativ
60	7	2.3	146	17	Q973F7	Q973f7 sulfolobus
61	7	2.3	158	2	Q9EWB3	Q9ewb3 streptomyc
62	7	2.3	158	16	Q8XSK1	Q8xsk1 ralestonia s
63	7	2.3	166	2	Q8GD28	Q8gd28 bordetella
64	7	2.3	171	16	Q06422	Q06422 mycobacteri
65	7	2.3	172	10	Q9SZQ8	Q9szq8 arabidopsis
66	7	2.3	173	5	Q22966	Q22966 caenorhabdi
67	7	2.3	173	10	Q64567	Q64567 arabidopsis
68	7	2.3	173	17	Q97ZK3	Q97zk3 sulfolobus
69	7	2.3	177	16	Q931Y0	Q931y0 staphylococ
70	7	2.3	177	16	Q8X7L2	Q8x7l2 escherichia
71	7	2.3	179	16	Q9RKC2	Q9rk2 streptomyc
72	7	2.3	189	17	Q8TWY8	Q8twy8 methanosarc
73	7	2.3	191	10	Q9XHM2	Q9xhm2 medicago tr
74	7	2.3	194	16	Q8EZ34	Q8ez34 leptospira
75	7	2.3	200	10	Q8VIT4	Q8vyt4 arabidopsis
76	7	2.3	202	6	Q9TUF4	Q9tuf4 propithecus
77	7	2.3	202	6	Q9TUF5	Q9tuf5 cercopithe
78	7	2.3	204	2	Q8KU14	Q8ku14 actinosyne
79	7	2.3	208	10	Q8O501	Q8o501 arabidopsis
80	7	2.3	208	10	Q40525	Q40525 nicotiana t
81	7	2.3	208	10	Q8H4Q9	Q8h4q9 oryza sativ
82	7	2.3	210	16	Q8DSW6	Q8dsw6 streptococc
83	7	2.3	211	16	Q9CDV2	Q9cdv2 lactococcus
84	7	2.3	214	2	Q32618	Q32618 helicobacte
85	7	2.3	214	6	Q62722	Q62722 oryctolagus
86	7	2.3	217	12	Q9YWN1	Q9ymn1 lymantria d
87	7	2.3	219	16	Q8KDL2	Q8kdl2 chlorobium
88	7	2.3	221	11	Q8C8M1	Q8c8m1 mus musculu
89	7	2.3	224	11	Q9R1M6	Q9rlm6 mus musculu

90 7 2.3 227 12 Q91BE2 Q91be2 spodoptera  
 91 7 2.3 229 16 Q8FR05 Q8fr05 corynebacte  
 92 7 2.3 230 2 Q8GIT0 Q8git0 synechococc  
 93 7 2.3 234 10 Q8H3H1 Q8h3h1 oryza sativ  
 94 7 2.3 242 16 Q8EGH5 Q8egh5 shewanella  
 95 7 2.3 242 17 Q9HPH6 Q9hph6 halobacteri  
 96 7 2.3 242 17 Q96XAL Q96xal sulfolobus  
 97 7 2.3 245 16 Q8FU79 Q8fu79 corynebacte  
 98 7 2.3 245 17 Q8PTV9 Q8ptv9 methanosarc  
 99 7 2.3 247 4 O7S192 O7s192 homo sapien  
 100 7 2.3 255 16 O50521 O50521 streptomyce

## ALIGNMENTS

## RESULT 1

Q9BX67 PRELIMINARY; PRT; 310 AA.  
 AC Q9BX67;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Junctional adhesion molecule 3 precursor (Junctional adhesion molecule-2) (Junctional adhesion molecule-3) (Hypothetical protein FLJ90288) (Hypothetical protein FLJ90828).  
 GN JAM-2 OR JAM3  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RA Cunningham S.A., Arrate M.P., Tran T.M.;  
 RT "Cloning of Human Junctional Adhesion Molecule 3."  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Aurand-Lions M.A., Johnson-leger C., Wong C., DuPasquier L.;  
 RT "Heterogeneity of endothelial junctions is reflected by differential expression and specific subcellular localization of the three JAM family members."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Aurand-Lions M.A., Johnson-leger C., Lamagna C., Ozaki H., Kita T.;  
 RT "Junctional adhesion molecules (JAMs) and interendothelial junctions."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Sachs U.J.H., Eva O., Berghoefer H., Santoso S.;  
 RT "Characterization of Junctional Adhesion Molecule-3 on Human Platelets: A New Member of Immunoglobulin Superfamily."  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Isozaki T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T., Suzuki Y., Nagai K., Sugano S., Ighii S., Kawai-Hio Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K., Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N., Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;  
 RT "NEO human cDNA sequencing project."  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF356518; AAK27221.1; -  
 DR EMBL; AJ344431; CAC69845.1; -  
 DR EMBL; AF448478; AAM20925.1; -  
 DR EMBL; AK074769; BAC11195.1; -  
 DR EMBL; AK075309; BAC11538.1; -  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003006; IG\_MHC.  
 PFam; PF00047; IG; 2.

DR PROSITE; PS50835; IG LIKE; 2.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 30 POTENTIAL.  
 SQ SEQUENCE 310 AA; 35020 MW; CE39ADF33EADAB9 CRC64;  
 Query Match 67.4%; Score 209; DB 4; Length 310;  
 Best Local Similarity 99.7%; Pred. No. 1.2e-211;  
 Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MALRRPRLRLCARLPDFFLLILFRGLIGAVNLKSSNRTVPVQEFESVELSCIITDSQT 60  
 DB 1 MALRRPRLRLCARLPDFFLLILFRGLIGAVNLKSSNRTVPVQEFESVELSCIITDSQT 60  
 QY 61 SDPRIEWKIQDEQTYVFFDNKIQGLAGRAEILGKTSIKIWNVTRDSALYRCVVAR 120  
 DB 61 SDPRIEWKIQDEQTYVFFDNKIQGLAGRAEILGKTSIKIWNVTRDSALYRCVVAR 120  
 QY 121 NDRKEIDBIVIELTVQVKPVPVCRPKAVPVGMATLHCQESGHPHPRHYSWYRNDVPL 180  
 DB 121 NDRKEIDBIVIELTVQVKPVPVCRPKAVPVGMATLHCQESGHPHPRHYSWYRNDVPL 180  
 QY 181 PTDSRANPRFRNSSHLNSETGTLVFTAVHKDDSGQYCIASNDAGSARCEQEMEVYDL 240  
 DB 181 PTDSRANPRFRNSSHLNSETGTLVFTAVHKDDSGQYCIASNDAGSARCEQEMEVYDL 240  
 QY 241 NIGGIIGVLVLAVALITLIGICAVRRGYFINNKQDGESYKPKGPDGVNYIRTDEEG 300  
 DB 241 NIGGIIGVLVLAVALITLIGICAVRRGYFINNKQDGESYKPKGPDGVNYIRTDEEG 300  
 QY 301 DFRHKSSFVI 310  
 DB 301 DFRHKSSFVI 310

## RESULT 2

Q8WWL8 PRELIMINARY; PRT; 355 AA.  
 AC Q8WWL8;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Junction adhesion molecule 3.  
 GN JAM3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hearn T.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Phillips H.M.;  
 RT "Narrowing the critical region within 11q24-qter for hypoplastic left heart and identification of a candidate gene, JAM3, expressed during cardiogenesis."  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ416101; CAC94776.1; -  
 DR Genew; HGNC:15532; JAM3.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003598; IG\_c2.  
 DR InterPro; IPR003006; IG\_MHC.  
 DR Pfam; PF00047; IG; 2.  
 DR SMART; SM00408; IGC2; 1.  
 DR PROSITE; PS50835; IG LIKE; 2.  
 KW Immunoglobulin domain.  
 FT CHAIN 76 355 JUNCTION ADHESION MOLECULE 3.  
 SQ SEQUENCE 355 AA; 39602 MW; 8B1577DEA7B1D4F8 CRC64;  
 Query Match 67.4%; Score 209; DB 4; Length 355;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-211;  
 Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MALRRPRLRLCARLPDFFLLFRGLIGAVNLKSSNRTPVVQEFESVELSCIITDSQT 60
DB 46 MALRRPRLRLCARLPDFFLLFRGLIGAVNLKSSNRTPVVQEFESVELSCIITDSQT 105
QY 61 SDPREWKKIQDEQTTVFFDNKIQDLAGRAEILGKTSLKINWVTRRDSALYRCEVARN 120
DB 106 SDPREWKKIQDEQTTVFFDNKIQDLAGRAEILGKTSLKINWVTRRDSALYRCEVARN 165
QY 121 NRKEIDEIVELTVQVKPVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPL 180
DB 166 NRKEIDEIVELTVQVKPVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPL 225
QY 181 PTDSRANPRFRNSSHLNSETGLVFTAVHKDDSGQYVCIAANDAGSARCEQEVEYDL 240
DB 226 PTDSRANPRFRNSSHLNSETGLVFTAVHKDDSGQYVCIAANDAGSARCEQEVEYDL 285
QY 241 NIGGIIGVVLVAVLALITLIGICAYRRGYFINNKQDGESYKNGKPGGVNVRTDERG 300
DB 286 NIGGIIGVVLVAVLALITLIGICAYRRGYFINNKQDGESYKNGKPGGVNVRTDERG 345
QY 301 DFRHKSSFFVI 310
DB 346 DFRHKSSFFVI 355

RESULT 3
Q96FL1
ID Q96FL1 PRELIMINARY; PRT; 309 AA.
AC Q96FL1;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RA EMBL; BC010690; AAH10690.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 2.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG LIKE; 2.
KW Hypothetical protein; Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 309 AA; 34917 MW; 50CSB1B7872E8DF3 CRC64;

Query Match 43.2%; Score 134; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 1.6e-132;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 ALRRPRLRLCARLPDFFLLFRGLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS 61
DB 1 ALRRPRLRLCARLPDFFLLFRGLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS 60
QY 62 DPRIEWKKIQDEQTTVFFDNKIQDLAGRAEILGKTSLKINWVTRRDSALYRCEVARN 121
DB 61 DPRIEWKKIQDEQTTVFFDNKIQDLAGRAEILGKTSLKINWVTRRDSALYRCEVARN 120
QY 122 DRKEIDEIVELTV 135
DB 121 DRKEIDEIVELTV 134

RESULT 4
Q9DBB7
ID Q9DBB7 PRELIMINARY; PRT; 310 AA.
AC Q9DBB7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE 110002N23Rik protein.
GN JCAM3 OR JCAM2 OR 1110002N23RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 408:685-690(2001).
DR EMBL; AK008187; BAB25519.1; -.
DR MGD; MGI:1933820; Jcam2.
DR MGD; MGI:1933825; Jcam3.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR003006; IG_MHC.
DR Pfam; PF00047; IG; 2.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50835; IG LIKE; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 310 AA; 34855 MW; C74884EABE234680 CRC64;

Query Match 10.3%; Score 32; DB 11; Length 310;
Best Local Similarity 100.0%; Pred. No. 6.8e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 PVVQEFESVELSCIITDSQTSDPRIEWKKIQD 72
DB 41 PVVQEFESVELSCIITDSQTSDPRIEWKKIQD 72

RESULT 5
Q9DIM9
ID Q9DIM9 PRELIMINARY; PRT; 310 AA.
AC Q9DIM9;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)
DE 110002N23Rik protein.
GN JCAM3 OR JCAM2 OR 1110002N23RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RC MEDLINE=21085660; PubMed=11217851;

```

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Asnburner M., Batalov S., Casavant T.,  
RA Fleschmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Blake J., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,  
RA Nordone P., Ring B., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK003326; BAB22715.1; -;  
DR MGD; MGI:1933820; Jcam3.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003598; IG\_c2.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; ig; 2.  
DR SMART; SM00408; IGC2; 1.  
DR PROSITE; PS0835; IG\_LIKE; 2.  
KW Immunoglobulin domain.  
SQ SEQUENCE 310 AA; 34819 MW; 6692BCAD68EA4B1D CRC64;  
Query Match 9.0%; Score 28; DB 11; Length 310;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 45 EFESVELSCIITDSQTSDPRIEWKKIQD 72  
Db |||||  
45 EFESVELSCIITDSQTSDPRIEWKKIQD 72  
RESULT 6  
Q9EPK4 PRELIMINARY; PRT; 310 AA.  
AC Q9EPK4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Functional adhesion molecule-2, JAM-2 (1110002N23Rik protein)  
DE (Junction cell adhesion molecule 3).  
GN JCAM3 OR JCAM2 OR JAM-2 OR 1110002N23RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=11036763;  
RA Aurand-Lions M.A., Duncan L., Du Pasquier L., Imhof B.A.;  
RT "Cloning of JAM-2 and JAM-3: an Emerging Junctional Adhesion Molecular  
RT Family";  
RL Curr. Top. Microbiol. Immunol. 251:91-98(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Embryo;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Asnburner M., Batalov S., Casavant H.,  
RA Fleschmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,  
RA Nordone P., Ring B., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RA "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Mesonephros;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RA 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AJ300304; CAC20704.1; -;  
DR EMBL; AK013156; BAB28683.1; -;  
DR EMBL; BC024357; AAH24357.1; -;  
DR EMBL; AK032833; BAC28049.1; -;  
DR MGD; MGI:1933820; Jcam2.  
DR MGD; MGI:1933825; Jcam3.  
DR InterPro; IPR007110; IG-like.  
DR InterPro; IPR003598; IG\_c2.  
DR InterPro; IPR003006; IG\_MHC.  
DR Pfam; PF00047; ig; 2.  
DR SMART; SM00408; IGC2; 1.  
DR PROSITE; PS0835; IG\_LIKE; 2.  
KW Immunoglobulin domain.  
SQ SEQUENCE 310 AA; 34837 MW; 4B92BCB51D0A4B0A CRC64;  
Query Match 9.0%; Score 28; DB 11; Length 310;  
Best Local Similarity 100.0%; Pred. No. 1.1e-20;  
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 45 EFESVELSCIITDSQTSDPRIEWKKIQD 72  
Db |||||  
45 EFESVELSCIITDSQTSDPRIEWKKIQD 72  
RESULT 7  
Q8BT59 PRELIMINARY; PRT; 64 AA.  
AC Q8BT59;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Junction cell adhesion molecule 3 (Fragment).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Body;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA "Analysis of the mouse transcriptome based on functional annotation of  
RA 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
DR EMBL; AK017692; BAC25526.1; -;

```
FT NON TER 1 1
SQ SEQUENCE 64 AA; 6996 MW; AF46BC30AAGDOC11 CRC64;

Query Match
Best Local Similarity 4.2%; Score 13; DB 11; Length 64;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 298 EGDPRHKSSFVI 310
Db 52 EGDPRHKSSFVI 64
|||||

RESULT 8
Q9TNT2 PRELIMINARY; PRT; 39 AA.
AC Q9TNT2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MHC class I HLA C antigen (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RX MEDLINE=99110146; PubMed=9894855;
RA van der Vlies S., Voorter C., van den Berg-Loonen E.;
RT "A reliable and efficient high resolution typing method for HLA C
using sequence-based typing.";
RL Tissue Antigens 52:558-568(1998).
DR EMBL; Y16418; CAA76206.1; -.
KW MHC.
FT NON TER 1
FT NON_TER 39
FT NON_TER 39
SQ SEQUENCE 39 AA; 3938 MW; 24613855AD783823 CRC64;

Query Match
Best Local Similarity 2.6%; Score 8; DB 7; Length 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 18 VLVVLAVL 25
|||||

RESULT 9
Q8HWG1 PRELIMINARY; PRT; 67 AA.
AC Q8HWG1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I antigen (Fragment).
GN HLA-CW.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-17 FROM N.A.
RA Delfino L., Morabito A., Ferrara G.B.;
RT "HLA-Cw*07012 sequence updating: from exon 6 to exon 8.";
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY162385; AAN84536.1; -.
DR EMBL; AY162382; AAN84536.1; JOINED.
DR EMBL; AY162383; AAN84536.1; JOINED.
DR EMBL; AY162384; AAN84536.1; JOINED.
FT NON TER 1
FT NON_TER 67
FT NON_TER 67
SQ SEQUENCE 67 AA; 6538 MW; 543B57F5F97D7FF1 CRC64;

Query Match
Best Local Similarity 2.6%; Score 8; DB 7; Length 67;
```

```
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 18 VLVVLAVL 25
|||||

RESULT 10
Q8I8S1 PRELIMINARY; PRT; 131 AA.
AC Q8I8S1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Odorant-binding protein G.38B.a.
OS Anopheles gambiae (African malaria mosquito).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=7165;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu P., Smith D.P.;
RT "Identification of distinct families of odorant-binding in the genome
of Anopheles gambiae.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY146730; AAO12090.1; -.
SQ SEQUENCE 131 AA; 14884 MW; FC7D52E9740E2A64 CRC64;

Query Match
Best Local Similarity 2.6%; Score 8; DB 5; Length 131;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 246 IGGVLVVL 253
Db 7 IGGVLVVL 14
|||||

RESULT 11
Q8C4V2 PRELIMINARY; PRT; 154 AA.
AC Q8C4V2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK080981; BAC38104.1; -.
KW Hypothetical protein.
SQ SEQUENCE 154 AA; 17098 MW; F032A4B8C5483EC6 CRC64;

Query Match
Best Local Similarity 2.6%; Score 8; DB 11; Length 154;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 224 DAGSARCE 231
Db 72 DAGSARCE 79
|||||

RESULT 12
```

```
Q8TQ5
ID Q8TQ5 PRELIMINARY; PRT; 231 AA.
AC Q8TQ5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cell surface lipoprotein.
GN MA1454.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nuebaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann K., DeAtrellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talanas J., Tirrell A., Ye W.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Guss A.M.,
RA Hederich R., Ingram-Smith C., Kuether H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,
RA Perry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity.";
RL Genome Res. 12:532-542(2002);
DR EMBL; AE010815; AAM04868.1; -.
KW Lipoprotein; Complete proteome.
SQ SEQUENCE 231 AA; 25437 MW; 7AAB21CE50B6500B CRC64;

Query Match 2.6%; Score 8; DB 17; Length 231;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 LWLAVLA 257
DB 23 LWLAVLA 30

RESULT 13
Q8ZSR8
ID Q8ZSR8 PRELIMINARY; PRT; 232 AA.
AC Q8ZSR8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein PAE3614.
GN PAE3614.
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009941; AAL65045.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 232 AA; 25942 MW; 39A5B8CC4403F903 CRC64;

Query Match 2.6%; Score 8; DB 17; Length 232;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 17 DFFLLLF 24
DB 2 DFFLLLF 9

RESULT 14
Q9LR49
ID Q9LR49 PRELIMINARY; PRT; 233 AA.
AC Q9LR49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE T25N20.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Johnson-Hopson C., Brooks S., Buehler E., Chao Q., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC T25N20 from chromosome
RT 1.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
RA Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
RA Ecker J.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005106; AAF79724.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 233 AA; 24829 MW; 28AD613B8085D58D CRC64;

Query Match 2.6%; Score 8; DB 10; Length 233;
```

Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 TDSQSDP 63  
Db 122 TDSQSDP 129

## RESULT 15

Q966C0 PRELIMINARY; PRT; 246 AA.  
AC Q966C0;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical protein Y45G12B.2b.  
GN Y45G12B.2  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Bradshaw H., Graves T., Fronick B.;  
RT "The sequence of C. elegans cosmid Y45G12B.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RT "Direct Submission";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006768; AAK70665.1; -;  
DR WormPep; Y45G12B.2b; CE28365.  
SQ SEQUENCE 246 AA; 27463 MW; CBEF536A113A2050 CRC64;

Query Match 2.6%; Score 8; DB 5; Length 246;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 253 LAVLALIT 260  
Db 199 LAVLALIT 206

## RESULT 16

Q29990 PRELIMINARY; PRT; 290 AA.  
AC Q29990;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE MHC class I lymphocyte antigen (Fragment).  
GN HLA-C.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89233295; PubMed=2714852;  
RA Pohl H., Kuon W., Tabaczewski P., Doerner C., Weiss E.H.;  
RT "Allelic variation in HLA-B and HLA-C sequences and the evolution of  
the HLA-B alleles.";

RL Immunogenetics 29:297-307(1989).  
CC -I- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
CC IMMUNE SYSTEM (BY SIMILARITY).  
CC -I- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
CC MICROGLOBULIN) (BY SIMILARITY).  
DR EMBL; M28207; AAS3259.1; -;  
DR HSSP; P30685; IASB.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR Pfam; PF00129; MHC\_I; 1.  
DR PRINTS; PR01638; MHCCLASSI.  
DR ProDom; PD000050; MHC\_I; 1.  
DR SMART; SM00407; IGcl; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
KW Glycoprotein; Transmembrane.  
FT NON TER 1  
SQ SEQUENCE 290 AA; 32253 MW; C17AB5E56E64A5A6 CRC64;

Query Match 2.6%; Score 8; DB 7; Length 290;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256  
Db 241 VLVVLAVL 248

## RESULT 17

Q98EQ9 PRELIMINARY; PRT; 294 AA.  
AC Q98EQ9;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Hypothetical protein ml14127.  
GN Mli4127.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003003; BAB50858.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 294 AA; 30557 MW; 7D20SD3C91E36A35 CRC64;

Query Match 2.6%; Score 8; DB 16; Length 294;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 251 VLVVLAVL 258  
Db 18 VLVVLAVL 25

## RESULT 18

Q9N4Y9 PRELIMINARY; PRT; 309 AA.  
AC Q9N4Y9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)

01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
Hypothetical 34.8 kDa protein.  
Y45G12B.2.  
Caenorhabditis elegans.  
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
Rhabditidae; Pelodierinae; Caenorhabditis.  
NCBI\_TaxID=6239;  
(1)  
SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Bradshaw H., Graves T., Fronick B.;  
RT "The sequence of C. elegans cosmid Y45G12B.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN (3)  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RT "Direct Submission.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC006768; AAF60574.1; -;  
DR WormPep; Y45G12B.2a; CE21934.  
DR InterPro; IPR001841; Znf\_ring.  
DR SMART; SM00184; RING\_1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 309 AA; 34815 MW; CF6DS930D1D0192D CRC64;  
Query Match 2.6%; Score 8; DB 5; Length 309;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 253 LAVLALIT 260  
Db 262 LAVLALIT 269  
RESULT 19  
Q9MY31  
ID Q9MY31 PRELIMINARY; PRT; 313 AA.  
AC Q9MY31;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Human leucocyte antigen C (Fragment).  
GN HLA-C.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99416947; PubMed=10488744;  
RA van der Vlies S.A., Voorster C.E., van den Berg-Loonen E.M.;  
RT "There is more to HLA -C than exons 2 and 3: sequencing exons 1, 4 and  
5.";  
RN Tissue Antigens 54:169-177 (1999).  
RL CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN) (BY SIMILARITY).  
DR EMBL; Y18533; CAB71800.1; -;  
DR EMBL; Y18534; CAB71800.1; JOINED.  
DR EMBL; Y18535; CAB71800.1; JOINED.  
DR EMBL; Y18536; CAB71800.1; JOINED.

HSSP; P30685; IAAE.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001039; MHC\_1.  
DR Pfam; PF00047; 19; 1.  
DR Pfam; PF00129; MHC\_I; 1.  
DR PRINTS; PR01638; MHC\_I\_1.  
DR PRODOM; PD000050; MHC\_I\_1.  
DR SMART; SM00407; Igcl\_1.  
DR PROSITE; PS50835; Ig\_LIKE; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 313 AA; 35441 MW; 2BFD5D879ECF238F CRC64;  
Query Match 2.6%; Score 8; DB 7; Length 313;  
Best Local Similarity 100.0%; Pred. No. 14;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 249 VLVVLAVL 256  
Db 292 VLVVLAVL 299  
RESULT 20  
Q9RP17  
ID Q9RP17 PRELIMINARY; PRT; 322 AA.  
AC Q9RP17;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE HYCD.  
GN HYCD.  
OS Desulfitobacterium dehalogenans.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Peptococcaceae;  
OC Desulfitobacterium.  
OX NCBI\_TaxID=36854;  
RN (1)  
RP SEQUENCE OF 1-99 FROM N.A.  
RX MEDLINE=20026799; PubMed=1059152;  
RA Smidt H., Song D., van der Oost J., de Vos W.M.;  
RT "Random transposition by Tn916 in Desulfitobacterium dehalogenans  
allows for isolation and characterization of halorespiration-deficient  
mutants.";  
RL J. Bacteriol. 181:6882-6888 (1999).  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Smidt H., Song D., van der Oost J., de Vos W.M.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.  
DR EMBL; AF157639; AAF13042.2; -;  
DR InterPro; IPR001694; Resp\_NADH\_dhl.  
DR Pfam; PF00146; NADHdh; 1.  
KW NAD; Oxidoreductase; Transmembrane.  
SQ SEQUENCE 322 AA; 34620 MW; 914CAF4BAC3F86AB CRC64;  
Query Match 2.6%; Score 8; DB 2; Length 322;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 254 AVLALITL 261  
Db 308 AVLALITL 315  
RESULT 21  
Q9C7E9  
ID Q9C7E9 PRELIMINARY; PRT; 336 AA.  
AC Q9C7E9;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)



DE RING zinc finger protein, putative.  
GN F13K9.14.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Alcafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,  
RA Dunn P., Egtu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldenmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,  
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
RT thaliana";  
RL Nature 408:816-820(2000).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AC069471; AAG51486.1; -;  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; Zf-C3HC4\_1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
SQ SEQUENCE 336 AA; 37711 MW; 1218C9624BE2PE4B CRC64;

Query Match 2.6%; Score 8; DB 10; Length 336;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 250 LVVLAVLA 257  
Db 28 LVVLAVLA 35  
|||||

RESULT 22  
Q9MXL5 PRELIMINARY; PRT; 338 AA.  
AC Q9MXL5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE MHC class I antigen (Fragment).  
GN PATR-C.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_taxid=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Blood;  
RX MEDLINE=20322475; PubMed=10866106;  
RA de Groot N.G., Otting N., Arguello R., Watkins D.I., Doxiadis G.G.M.,  
RA Madrigal J.A., Bontrop R.E.;  
RT "Major histocompatibility complex class I diversity in a West African  
RT chimpanzee population: implications for HIV research.";  
RL Immunogenetics 51:398-409(2000).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
CC IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN) (BY SIMILARITY).  
DR EMBL; AF165373; AAF72772.1; -;  
DR HSSP; P30685; IAGE.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR0129; MHC I; 1.  
DR PRODOM; PD000050; MHC I; 1.  
DR SMART; SM00407; IGC1\_1.  
DR PROSITE; PS50835; IG LIKE; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 338  
SQ SEQUENCE 338 AA; 38385 MW; 2DC19A1458352DEF CRC64;

Query Match 2.6%; Score 8; DB 7; Length 338;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 249 VLVVLAVL 256  
Db 317 VLVVLAVL 324  
|||||

RESULT 23  
Q9MWJ9 PRELIMINARY; PRT; 348 AA.  
AC Q9MWJ9;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE MHC class I heavy chain antigen (Fragment).  
GN GOGO-C.  
OS Gorilla gorilla (gorilla).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.  
OX NCBI\_taxid=9593;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21142798; PubMed=11207289;  
RA Urwater J.A., Hickman H., Dzuris J.L., Prillman K., Allen T.M.,  
RA Schwartz K.J., Lorentzen D., Shuffelebotham C., Collins E.J.,  
RA Neiffer D.L., Raphael B., Hildebrand W., Sette A., Watkins D.I.;  
RT "Gorillas with Spondyloarthropathies Express an MHC Class I Molecule  
RT with Only Limited Sequence Similarity to HLA-B\*27 that Binds Peptides  
RT with Arginine at P2.";  
RL J. Immunol. 166:3334-3344(2001).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
CC IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-

CC MICROGLOBULIN) (BY SIMILARITY).  
DR EMBL; AF157411; AAF80354.1; -;  
DR HSSP; P03989; IHSA.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig\_cl.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR0129; MHC I; 1.  
DR PRODOM; PD000050; MHC I; 1.  
DR SMART; SM00407; IGC1\_1.  
DR PROSITE; PS50835; IG LIKE; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 348  
SQ SEQUENCE 348 AA; 38753 MW; 260F31EBCDBP620F CRC64;

Query Match 2.6%; Score 8; DB 7; Length 348;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      249 VLVVLAVL 256
DB      307 VLVVLAVL 314
|||||
|||||

RESULT 24
Q9PR46      PRELIMINARY;      PRT;      348 AA.
AC Q9PR46;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Type I restriction enzyme M protein (Fragment).
GN HSDM-1 OR UO98.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Setovar 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Casseil G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RT Nature 407:757-762 (2000).
DR EMBL; AE002110; RAF30504.1; -.
DR InterPro; IPR004546; HsdM.
DR InterPro; IPR003665; Methylase M.
DR InterPro; IPR002296; N12N6_mtfase.
DR InterPro; IPR003356; N6_DNA_Mtase.
DR InterPro; IPR002052; N6_Mtase.
DR InterPro; IPR000051; SAM_bind.
DR Pfam; PF02506; Methylase M; 1.
DR Pfam; PF02384; N6_Mtase; 1.
DR PRINTS; PR00507; N12N6MTFRASE.
DR TIGRFAMs; TIGR00497; hsdM; 1.
DR PROSITE; PS00092; N6_MTASE; 1.
KW Complete proteome.
SQ SEQUENCE 348 AA; 39547 MW; 8D776CB13681DB62 CRC64;

Query Match      2.6%; Score 8; DB 16; Length 348;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      123 RKEIDEIV 130
DB      336 RKEIDEIV 343
|||||
|||||

RESULT 25
Q9SHC2      PRELIMINARY;      PRT;      363 AA.
AC Q9SHC2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Major histocompatibility complex, class I, C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye, and Retinal pigment epithelium;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC -!- IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; BC007814; AAH07814.1; -.

```

```

DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; MHC; Transmembrane.
SQ SEQUENCE 363 AA; 40262 MW; AA851E80336F24F8 CRC64;

Query Match      2.6%; Score 8; DB 7; Length 363;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      249 VLVVLAVL 256
DB      314 VLVVLAVL 321
|||||
|||||

RESULT 26
Q9XRX8      PRELIMINARY;      PRT;      365 AA.
AC Q9XRX8;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE MHC class I antigen (Fragment).
GN POPY C.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99367395; PubMed=10436180;
RA Adams E.J., Thomson G., Parham P.;
RT "Evidence for an HLA-C-like locus in the orangutan Pongo pygmaeus.";
RL Immunogenetics 49:865-871 (1999).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC -!- IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF118897; AAD28437.1; -.
DR HSPF; P30685; IAPF.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
FT NON TER 1
SQ SEQUENCE 365 AA; 40722 MW; 4C9BB3A18B860F0B CRC64;

Query Match      2.6%; Score 8; DB 7; Length 365;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      249 VLVVLAVL 256
DB      316 VLVVLAVL 323
|||||
|||||

RESULT 27
Q9XRX7

```

Q9RXK7 PRELIMINARY; PRT; 365 AA.  
AC Q9RXK7;  
DT 01-NOV-1999 (TREMELrel. 12, Created)  
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE MHC class I antigen (Fragment).  
POPY C.  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo..  
OX NCBI\_TaxID=9600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99367395; PubMed=10436180;  
RA Adams E.J., Thomson G., Parham P.;  
RT "Evidence for an HLA-C-like locus in the orangutan Pongo pygmaeus.";  
RL Immunogenetics 49:865-871(1999).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN) (BY SIMILARITY).  
CC EMBL; AF118898; AAD28438.1; -.  
DR HSSP; P30685; 1A9E.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig cl.  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR01638; MHCCLASSI.  
DR ProDom; PD000050; MHC\_I; 1.  
DR SMART; SM00407; IGcl; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
DR PROSITE; PS00290; IG MHC; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 365 AA; 40613 MW; D3C078CFE226EB1D CRC64;  
Query Match 2.6%; Score 8; DB 7; Length 365;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 249 VLVVLAVL 256  
Db 316 VLVVLAVL 323  
RESULT 28  
Q9MXG4 PRELIMINARY; PRT; 365 AA.  
AC Q9MXG4;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DT 01-WAR-2003 (TREMELrel. 23, Last annotation update)  
DE MHC class I antigen (Fragment).  
GN PATR-C.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=2032476; PubMed=10866107;  
RA Adams E.J., Cooper S., Thomson G., Parham P.;  
RT "Common Chimpanzees have Greater Diversity than Humans at Two of the  
Three Highly Polymorphic MHC Class I Genes.";  
RL Immunogenetics 51:410-424(2000).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN) (BY SIMILARITY).  
CC EMBL; AF179672; AAF76974.1; -.  
DR HSSP; P30685; 1A9E.

DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig cl.  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR01638; MHCCLASSI.  
DR ProDom; PD000050; MHC\_I; 1.  
DR SMART; SM00407; IGcl; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
KW Glycoprotein; Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 365 AA; 40842 MW; 51247CE97E8866E5 CRC64;  
Query Match 2.6%; Score 8; DB 7; Length 365;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 249 VLVVLAVL 256  
Db 316 VLVVLAVL 323  
RESULT 29  
O46684 PRELIMINARY; PRT; 366 AA.  
AC O46684;  
DT 01-JUN-1998 (TREMELrel. 06, Created)  
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)  
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
DE PATR-C\*0901.  
GN PATR-C\*09012 OR PAPA-C\*03011 OR PAPA-C\*03012 OR PATR-C\*09011.  
OS Pan troglodytes (Chimpanzee), and  
Pan paniscus (Pygmy Chimpanzee) (Bonobo).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9598, 9597;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9809755; PubMed=9435339;  
RA Cooper S.L., Adams E.J., Wells R.S., Walker C.M., Parham P.;  
RT "A major histocompatibility complex class I allele shared by two  
species of chimpanzee.";  
RL Immunogenetics 47:212-217(1998).  
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  
IMMUNE SYSTEM (BY SIMILARITY).  
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  
MICROGLOBULIN) (BY SIMILARITY).  
CC EMBL; AJ001976; CAA05124.1; -.  
DR EMBL; AJ001975; CAA05123.1; -.  
DR EMBL; AJ001973; CAA05121.1; -.  
DR EMBL; AJ001974; CAA05122.1; -.  
DR HSSP; P30685; 1A9E.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003597; Ig cl.  
DR InterPro; IPR003006; Ig MHC.  
DR InterPro; IPR001039; MHC\_I.  
DR Pfam; PF00047; Ig; 1.  
DR PRINTS; PR01638; MHCCLASSI.  
DR ProDom; PD000050; MHC\_I; 1.  
DR SMART; SM00407; IGcl; 1.  
DR PROSITE; PS00835; IG LIKE; 1.  
KW Glycoprotein; Transmembrane.  
SQ SEQUENCE 366 AA; 40965 MW; 42D959BDB8CC92C01 CRC64;  
Query Match 2.6%; Score 8; DB 6; Length 366;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 249 VLVVLAVL 256  
Db 317 VLVVLAVL 324

```
RESULT 30
Q8MH0 PRELIMINARY; PRT; 366 AA.
AC Q8MH0;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE MHC class I antigen.
GN POPY-C.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072192; PubMed=12077248;
RA Guethlein L.A., Flodin L.R., Adams E.J., Parham P.;
RT "NK Cell Receptors of the Orangutan (Pongo pygmaeus): A Pivotal
RT Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors
RT with MHC-C.";
RL J. Immunol. 169:220-229 (2002).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
CC EMBL; AF470377; AAM78477.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40672 MW; 05DC89FAD0DB831A CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 317 VLVVLAVL 324

RESULT 31
Q9TQ9 PRELIMINARY; PRT; 366 AA.
AC Q9TQ9;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE HLA class I antigen precursor.
GN HLA-Cw*070X.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99299762; PubMed=10372547;
RA Baurain J.F., Coulie P.G.;
RT "Correction of HLA-Cw*0501 and identification of HLA-Cw*0711.";
RL Tissue Antigens 53:510-512 (1999).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
```

```
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AJ010749; CAA09341.1; -.
DR HSP; P30460; IAGD.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 366 HLA CLASS I ANTIGEN.
SQ SEQUENCE 366 AA; 40789 MW; 4073C3A023E55DOC CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 317 VLVVLAVL 324

RESULT 32
Q29652 PRELIMINARY; PRT; 366 AA.
AC Q29652;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE HLA-Cw7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RA Coulie P.G.;
RT "Identification of a new HLA-Cw7 allele.";
RL Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
CC EMBL; U09853; AAA50217.1; -.
DR HSP; P30460; IAGD.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40847 MW; 63B21ECD158049B6 CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 317 VLVVLAVL 324
```

```
RESULT 33
Q95463
ID Q95463 PRELIMINARY; PRT; 366 AA.
AC Q95463;
DT 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE HLA-C protein precursor.
GN HLA-CW*07GB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97161041; PubMed=9008313;
RA Vilches C., Bunce M., Sanz L., de Pablo R., Puente S., Kreisler M.;
RT "Molecular cloning of two new HLA-C alleles: Cw*1801 and Cw*0706."
RL Tissue Antigens 48:698-702(1996).
RC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; X97321; CA65986.1; -.
DR HSP; P30685; I49E.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 366 POTENTIAL.
SQ SEQUENCE 366 AA; 40735 MW; 0B95005400280609 CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324

RESULT 34
Q95463
ID Q95463 PRELIMINARY; PRT; 366 AA.
AC Q95463;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MHC class I antigen.
GN POPY-C.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072192; PubMed=12077248;
RA Guethlein L.A., Flodin L.R., Adams E.J., Parham P.;
RT "NK Cell Receptors of the Orangutan (Pongo pygmaeus): A Pivotal
RT Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors
RT with MHC-C."
RL J. Immunol. 169:220-229(2002).
RC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF470379; AAM78479.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40890 MW; 749AEF8A8624D22D CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324
```

```
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF470379; AAM78479.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40853 MW; 3132BC758471B1EC CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324

RESULT 35
Q95463
ID Q95463 PRELIMINARY; PRT; 366 AA.
AC Q95463;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE MHC class I antigen.
GN POPY-C.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072192; PubMed=12077248;
RA Guethlein L.A., Flodin L.R., Adams E.J., Parham P.;
RT "NK Cell Receptors of the Orangutan (Pongo pygmaeus): A Pivotal
RT Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors
RT with MHC-C."
RL J. Immunol. 169:220-229(2002).
RC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF470379; AAM78479.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; IGC1; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
DR Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40890 MW; 749AEF8A8624D22D CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324
```

```

RESULT 36
Q8MHG9 PRELIMINARY; PRT; 366 AA.
ID Q8MHG9
AC Q8MHG9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE MHC class I antigen.
CN POPY-C.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22072192; PubMed=12077248;
RA Guthlein L.A., Flodin L.R., Adams E.J., Parham P.;
RT "NK Cell Receptors of the Orangutan (Pongo pygmaeus): A Pivotal
RT Species for Tracking the Coevolution of Killer Cell Ig-Like Receptors
RT with MHC-C."
RL J. Immunol. 169:220-229(2002).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC -!- IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AF470378; AAM78478.1; -.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I.1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I.1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40744 MW; F4513EAD84036097 CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324
|||||

RESULT 37
O19617 PRELIMINARY; PRT; 366 AA.
ID O19617
AC O19617;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HLA class I antigen (HLA-C protein).
CN HLA-C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97321045; PubMed=9177776;
RA Mizuki N., Ando H., Kimura M., Ohno S., Miyata S., Yamazaki M.,
RA Tashiro H., Watanabe K., Ono A., Taguchi S., Sugawara C., Fukuzumi Y.,
RA Okumura K., Goto K., Ishihara M., Nakamura S., Yonemoto J.,
RA Kikuti Y.Y., Shina T., Chen L., Ando A., Ikemura T., Inoko H.;
RT "Nucleotide sequence analysis of the HLA class I region spanning the
RT 237- kb segment around the HLA-B and -C genes."
RL Genomics 42:55-66(1997).

```

```

[2]
RN RP SEQUENCE FROM N.A.
RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Shina S., Tamiya G., Oka A., Inoko H.;
RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; D83957; BAA22206.1; -.
DR HSSP; P30885; IAPe.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I.1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I.1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Glycoprotein; Transmembrane.
SQ SEQUENCE 366 AA; 40710 MW; 0B9518E6AA77A609 CRC64;

Query Match 2.6%; Score 8; DB 7; Length 366;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 VLVVLAVL 256
DB 317 VLVVLAVL 324
|||||

RESULT 38
O78083 PRELIMINARY; PRT; 366 AA.
ID O78083
AC O78083;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE HLA-Cw*0702.
GN HLA-Cw*0702.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98099755; PubMed=9435339;
RA Cooper S.L., Adams E.J., Wells R.S., Walker C.M., Parham P.;
RT "A major histocompatibility complex class I allele shared by two
RT species of chimpanzee."
RL Immunogenetics 47:212-217(1998).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
CC IMMUNE SYSTEM (BY SIMILARITY).
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN) (BY SIMILARITY).
DR EMBL; AJ001977; CAA05125.1; -.
DR HSSP; P30885; IAPe.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003597; Ig cl.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00129; MHC_I.1.
DR PRINTS; PR01638; MHCCLASSI.
DR ProDom; PD000050; MHC_I.1.
DR SMART; SM00407; IGcl; 1.

```

RESULT 40  
Q95603  
ID Q95603

<p> [1]  SEQUENCE FROM N.A.  RP  RX MEDLINE=20322476; PubMed=10866106;  RA Adams E.J., Cooper S., Thomson G., Parham P.;  RT "Common Chimpanzees have Greater Diversity than Humans at Two of the  RT Three Highly Polymorphic MHC Class I Genes";  RL Immunogenetics 51:410-424(2000).  [2]  SEQUENCE FROM N.A.  RP  RC TISSUE=Blood;  RX MEDLINE=20322475; PubMed=10866106;  RA de Groot N.G., Oetting N., Arguello R., Watkins D.I., Doxiadis G.G.M.,  RA Madrigal J.A., Ontrop R.E.;  RT "Major histocompatibility complex class I diversity in a West African  RT chimpanzee population: implications for HIV research.";  RL Immunogenetics 51:398-409(2000).  CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE  CC IMMUNE SYSTEM (BY SIMILARITY).  CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-  CC MICROGLOBULIN) (BY SIMILARITY).  CC EMBL; AF179671; AAF76973.1; -;  CC EMBL; AF165366; AAF72765.1; -;  CC HSSP; P30685; 1A9E.  CC InterPro; IPR007110; Ig-like.  CC InterPro; IPR003006; Ig MHC.  CC InterPro; IPR001039; MHC_I.  CC Pfam; PF00047; Ig; 1.  CC Pfam; PF00129; MHC I; 1.  CC PRINTS; PR01638; MHCCLASSI.  CC ProDom; PD000050; MHC_I; 1.  CC PROSITE; PS00835; IG_LIKE; 1.  CC Glycoprotein; MHC; Transmembrane.  CC KW  CC SEQUENCE 366 AA; 40965 MW; 42D959BD8CC92C01 CRC64; </p>	<p> Query Match 2.6%; Score 8; DB 7; Length 366;  Best Local Similarity 100.0%; Pred. No. 17;  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps </p>
<p> QY 249 VLVVLAVL 256       DB 317 VLVVLAVL 324 </p>	
<p> RESULT 43  Q8CWZ7 PRELIMINARY; PRT; 449 AA.  AC Q8CWZ7;  DT 01-MAR-2003 (TrEMBLrel. 23, Created)  DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  DE Putative cation efflux pump (Multidrug resistance protein).  GN SMU 71.  OS Streptococcus mutans.  OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  OC Streptococcus.  OX NCBI_TaxID=1309;  [1]  SEQUENCE FROM N.A.  RC STRAIN=UA159 / ATCC 700610 / Serotype C;  RX MEDLINE=22295063; PubMed=12397186;  RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,  RA Carson M.B., Nijmeh C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,  RA Li S., Zhu H., Prayau P., Lai H., White J., Roe B.A., Ferretti J.J.;  RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental  RT pathogen.";  RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).  DR EMBL; AE014859; AAAS7857.1; -;  CC Complete proteome.  CC SEQUENCE 449 AA; 48324 MW; C5304BD28AB09EE5 CRC64; </p>	<p> Query Match 2.6%; Score 8; DB 16; Length 449;  Best Local Similarity 100.0%; Pred. No. 20;  Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps </p>



```

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 492 AA; 51279 MW; E3D46E40C287D57B CRC64;

Query Match      2.6%; Score 8; DB 16; Length 492;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 242 IGGIIGGV 249
Db 244 IGGIIGGV 251

RESULT 46
Q8EK37 PRELIMINARY; PRT; 659 AA.
AC Q8EK37;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cytochrome c-type biogenesis protein CcmF.
GN CCMF-1 OR S00266.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beaman M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.P.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vanatrevan J., Weidman J., Impraim M., Lee K., Barry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Otterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealon K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015474; AAN53351.1; -.
DR TIGR; S00266; -.
KW Complete proteome.
SQ SEQUENCE 659 AA; 72256 MW; 5FDE5B6D4BE76662 CRC64;

Query Match      2.6%; Score 8; DB 16; Length 659;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 252 VLAVALAI 259
Db 125 VLAVALAI 132

RESULT 47
Q9CIF4 PRELIMINARY; PRT; 784 AA.
AC Q9CIF4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Carbon starvation protein.
GN CSTA OR L10409.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Lactobacilliales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarne K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus

```

```

Qy 202 GTLVFTAV 209
Db 369 GTLVFTAV 376

RESULT 44
Q8GSD2 PRELIMINARY; PRT; 452 AA.
AC Q8GSD2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical transmembrane protein in upf0013.
GN BL1082.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwielen M.-C., Desiere P., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).
DR EMBL; AE014731; AAN24890.1; -.
KW Hypothetical protein; Transmembrane; Complete proteome.
SQ SEQUENCE 452 AA; 47921 MW; 5D9733BB97B22873 CRC64;

Query Match      2.6%; Score 8; DB 16; Length 452;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 249 VLVVLAVL 256
Db 335 VLVVLAVL 342

RESULT 45
Q8Z1U9 PRELIMINARY; PRT; 492 AA.
AC Q8Z1U9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative membrane protein.
GN STV4415.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagsis K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrrell B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.";
RL Nature 413:848-852 (2001).
DR EMBL; AL627282; CAD09203.1; -.
DR InterPro; IPR003352; PTS_EIIC.
DR Pfam; PF02378; PTS_EIIC; 1.

```

RT lactis ssp. lactis IL1403.";  
RL Genome Res. 11:731-753(2001).  
DR EMBL; AE006277; AAK04507.1; -.  
DR InterPro; IPR003706; CstA.  
DR Pfam; PF02554; CstA; 1.  
KW Complete proteome.  
SQ SEQUENCE 784 AA; 85323 MW; 38F7B2251765158C CRC64;

Query Match 2.6%; Score 8; DB 16; Length 784;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 252 VLAVLALI 259  
|||  
Db 223 VLAVLALI 230

## RESULT 48

Q9W6E1 PRELIMINARY; PRT; 1290 AA.  
AC Q9W6E1;  
DT 01-NOV-1999 (TReMBLrel. 12, Created)  
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
DT 01-WAR-2003 (TReMBLrel. 23, Last annotation update)  
DE Neurocan core protein.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1] -  
RP SEQUENCE FROM N.A.

RA MEDLINE=20309833; PubMed=10851024;  
RA Li H., Leung T.C., Hoffman S., Balsamo J., Lillien J.;  
RT "Coordinate Regulation of Cadherin and Integrin Function by the  
RT Chondroitin Sulfate Proteoglycan Neurocan.";  
RL J. Cell Biol. 149:1275-1288(2000).  
DR EMBL; AF116856; AAD24546.2; -.  
DR HSP; P08709; 1BF9.

DR InterPro; IPR000152; Asx hydroxyl.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR003599; Ig\_  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR001304; Lectin\_C.  
DR InterPro; IPR000538; Link.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi; 1.  
DR Pfam; PF00193; Xlink; 2.  
DR PRINTS; PR00010; EGFLOOD.  
DR PRINTS; PR01265; LINKMODULE.  
DR PRODOM; PD000918; Link; 2.  
DR SMART; SM00032; CCP; 1.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00445; LINK; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE LECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS00835; IG\_LINK; 1.  
DR PROSITE; PS01241; LINK; 2.  
KW EGF-like domain.

SQ SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;

Query Match 2.6%; Score 8; DB 13; Length 1290;  
Best Local Similarity 100.0%; Pred. No. 56;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 112 LYRCEVVA 119  
|||||  
Db 137 LYRCEVVA 144

## RESULT 49

Q8NI85 PRELIMINARY; PRT; 67 AA.  
AC Q8NI85;  
DT 01-OCT-2002 (TReMBLrel. 22, Created)  
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE OK/SW-CL.4.  
DE OK/SW-CL.4.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1] -  
RP SEQUENCE FROM N.A.  
RA Shichijo S., Itoh K.;  
RT "Identification of immuno-peptidmics that recognized by tumor-reactive  
RT CTL generated from TIL of colon cancer patients.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB062435; BAB93498.1; -.  
SQ SEQUENCE 67 AA; 7517 MW; 3BCD63D8E72609C6 CRC64;

Query Match 2.3%; Score 7; DB 4; Length 67;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 FFLLLLF 24  
|||||  
Db 52 FFLLLLF 58

## RESULT 50

Q99GX8 PRELIMINARY; PRT; 85 AA.  
AC Q99GX8;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)  
DE ORF72 (Hypothetical 10.0 kDa protein).  
OS Helicoverpa zea single nucleocapsid nucleopolyhedrovirus,  
OS Helicoverpa armigera nucleopolyhedrovirus G4, and  
OS Helicoverpa armigera nuclear polyhedrosis virus.  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;  
OC Nucleopolyhedrovirus.  
OX NCBI\_TaxID=10468, 148363, 51313;  
RN [1] -  
RP SEQUENCE FROM N.A.  
RC SPECIES=Helicoverpa zea single nucleocapsid nucleopolyhedrovirus;  
RA Chen X., Zhang W.-J., Wong J., Chun G., Lu A., McCutchen B.F.,  
RA Presnall J.K., Herrmann R., Dolan M., Tingey S., Hu Z.-H., Vlak J.M.;  
RT "Genome sequence analysis of Helicoverpa zea single nucleocapsid  
RT nucleopolyhedrovirus.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [2] -  
RP SEQUENCE FROM N.A.  
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
RA Deng F., Chen X., Vlak J.M., Arif B.M., Hu Z.;  
RT "Sequence analysis of the gp37 gene of Heliothis armigera single-  
RT nucleocapsid nucleopolyhedrovirus.";  
RL Zhongguo Bingduxue 15:35-42(2000).  
RN [3] -  
RP SEQUENCE FROM N.A.  
RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;

RA Wang H., Hu Z., Sun X., Vlask J.M., Chen X.;  
 RT "Sequence analysis of the iap3 gene of Heliothis armigera single-  
 RT nucleocapsid nucleopolyhedrovirus";  
 RN Zhongguo Bingduxue 15:43-49(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RX MEDLINE=21078302; PubMed=11210934;  
 RA Wang H., Chen X., Wang H., Arif B.M., Vlask J.M., Hu Z.;  
 RT "Nucleotide sequence and transcriptional analysis of a putative basic  
 RT DNA-binding protein of Helicoverpa armigera polyhedrovirus";  
 RL Virus Genes 22:113-120(2001).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RX MEDLINE=21064569; PubMed=11125177;  
 RA Chen X., Ijkel W.F., Tarchini R., Sun X., Sandbrink H., Wang H.,  
 RA Peters S., Zuidema D., Lankhorst R.K., Vlask J.M., Hu Z.;  
 RT "The sequence of the Helicoverpa armigera single-nucleocapsid  
 RT nucleopolyhedrovirus genome";  
 RL J. Gen. Virol. 82:241-257(2001).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nucleopolyhedrovirus G4;  
 RX PubMed=12050807;  
 RA Zhang C.X., Wu J.C.;  
 RT "Genome structure and the p10 gene of the Helicoverpa armigera  
 RT nucleopolyhedrovirus";  
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=C1;  
 RX PubMed=12050807;  
 RA Zhang C.X., Wu J.C.;  
 RT "Genome structure and the p10 gene of the Helicoverpa armigera  
 RT nucleopolyhedrovirus";  
 RL Sheng Wu Hua Hsueh Yu Sheng Wu Wu Li Hsueh Pao 33:179-184(2001).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Helicoverpa armigera nuclear polyhedrosis virus; STRAIN=C1;  
 RX PubMed=12050807;  
 RA Zhang C.X., Jin W.R.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF334030; AAL56078.1; -;  
 DR EMBL; AF271059; AAG53813.1; -;  
 DR EMBL; AF303045; AAK96317.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 85 AA; 9958 MW; 6560D8FAA06A6787 CRC64;

Query Match 2.3%; Score 7; DB 12; Length 85;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 FFLLLLF 24  
 |||||  
 Db 28 FFLLLLF 34

Search completed: December 15, 2003, 15:00:24  
 Job time : 46 secs

**THIS PAGE BLANK (uspto)**